Monatshefte für Mathematik (2020) 192:225–247
https://doi.org/10.1007/s00605-020-01393-9

On multiplicative independence of rational function iterates

Marley Young[1]

Received: 23 September 2017 / Accepted: 15 February 2020 / Published online: 21 February 2020

Abstract
We give lower bounds for the degree of multiplicative combinations of iterates of rational functions (with certain exceptions) over a general field, establishing the multiplicative independence of said iterates. This leads to a generalisation of Gao's method for constructing elements in the finite field $\mathbb{F}_{q^n}$ whose orders are larger than any polynomial in n when n becomes large. Additionally, we discuss the finiteness of polynomials which translate a given finite set of polynomials to become multiplicatively dependent.

Keywords Iteration · Multiplicative dependence · Rational function

Mathematics Subject Classification 11R18 · 39B12 · 12E99 · 37F10

1 Introduction and main results

We say that n non-zero elements $a_1, \ldots, a_n$ of a ring are multiplicatively independent if, for integers $k_1, \ldots, k_n$, we have that $a_1^{k_1} \ldots a_n^{k_n} = 1$ if and only if $k_1 = \cdots = k_n = 0$. Otherwise we say they are multiplicatively dependent. Multiplicative independence, especially of values of polynomials and rational functions, is being increasingly studied. In [4], Bombieri, Masser and Zannier initiate study of the intersection of algebraic curves with proper algebraic subgroups of the multiplicative group $\mathbb{G}_m^n$. It turns out (see [3, Corollary 3.2.15]) that each such subgroup of $\mathbb{G}_m^n$ is defined by finitely many equations of the form $X_1^{k_1} \ldots X_n^{k_n} = 1$, where $k_1, \ldots, k_n$ are integers, not all zero. As such, [4], which leads into the area of "unlikely intersections", really concerns the multiplicative dependence of points on curves.

Communicated by Umberto Zannier.

✉ Marley Young
mjy28@cam.ac.uk

[1] Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge CB3 0WB, UK

More recently, we see multiplicative independence being studied in the context of arithmetic dynamics. In [18], it is shown that under fairly natural conditions on rational functions $f_1, \ldots, f_s$ over a number field $\mathbb{K}$, the values $f_1(\alpha), \ldots, f_s(\alpha)$ are multiplicatively independent for all but finitely many $\alpha \in \mathbb{K}^{\mathrm{ab}}$, where $\mathbb{K}^{\mathrm{ab}}$ is the maximal abelian extension of $\mathbb{K}$. This leads to results on multiplicative dependence in the orbits of a univariate polynomial dynamical system.

Clearly, to study the multiplicative independence of elements in the orbits of polynomials or rational functions, it is necessary to know when the given functions are multiplicatively dependent, as in this case all their values must be multiplicatively dependent. We study this problem in the context of iterates of rational functions over a field.

Throughout the paper, $\mathbb{F}$ will denote a field of characteristic p (zero or prime), and $f \in \mathbb{F}(X)$ a non-constant rational function in lowest terms over $\mathbb{F}$. That is, $f = g/h$ with $d := \deg f = \max\{\deg g, \deg h\} \geq 1$. Being in "lowest terms" means $\gcd(g, h) = 1$, or equivalently, g and h share no roots in any extension field of $\mathbb{F}$. As such, when referring to zeros and poles of a rational function, we mean roots of its numerator and denominator respectively in an algebraic closure $\overline{\mathbb{F}}$ of $\mathbb{F}$. We recursively define the iterates of f by

$$f^{(0)}(X) = X, \quad \text{and} \quad f^{(k)} = f \circ f^{(k-1)} \text{ for } k \geq 1.$$

In [10], Gao considers the multiplicative independence of polynomials over a finite field $\mathbb{F}_q$, where q is a prime power, proving that if $f \in \mathbb{F}_q[X]$ is not a monomial or certain binomial, then the iterates $f^{(1)}, \ldots, f^{(n)}$ are multiplicatively independent for $n \geq 1$. Gao uses this fact to give a method for constructing elements of "high order" in $\mathbb{F}_{q^n}$ when q is fixed. That is, elements with order larger than any polynomial in n when n is large. In particular, if we define $\bar{n} = q^{\lceil \log_q n \rceil}$, and $g \in \mathbb{F}_q[X]$ is not a monomial or certain binomial, then any root of an irreducible factor of degree n of $X^{\bar{n}} - g(X)$ is an element in $\mathbb{F}_{q^n}$ of order at least

$$n^{\frac{\log_q n}{4\log_q(2\log_q n)} - \frac{1}{2}}.$$

Sharper analysis of the same method by Popovych in [19] improves the lower bound on the order to

$$\binom{n+t-1}{t} \prod_{i=0}^{t-1} \frac{1}{d^i},$$

where $d = \lceil 2\log_q n \rceil$ and $t = \lfloor \log_d n \rfloor$.

In the case of rational functions over a general field, we also have multiplicative independence of iterates, up to a few exceptional cases. We remark (see Lemma 2.8) that these exceptions are precisely the rational functions which, under iteration, eventually become a monomial. For example, if $f^{(n)}(X) = X^k$, then $f^{(n)}(X)$ and $f^{(2n)}(X) = X^{k^2}$ are multiplicatively dependent. Note also that the cases of zero and

positive characteristic are different. One distinction, of course, is the existence of inseparable maps in fields of positive characteristic. We see in Lemma 2.7, that this corresponds to a difference in which rational functions have an iterate which is a polynomial, let alone a monomial. Moreover, especially in the polynomial case, positive characteristic allows terms in iterates to vanish which would otherwise prevent them from becoming monomials.

Theorem 1.1 *Suppose that $f = g/h \in \mathbb{F}(X)$ has degree $d \geq 2$, and is not a monomial of the form $aX^{\pm d}$, nor of the form $L(X^{p^\ell})$, where $L \in \mathbb{F}(X)$ has degree 1 and ℓ is a positive integer. Let $n \geq 1$, and write*

$$\Psi(n) = \min_{\substack{k_1,\dots,k_n \in \mathbb{Z} \\ k_n \neq 0}} \left(\deg \left(\left(f^{(1)} \right)^{k_1} \dots \left(f^{(n)} \right)^{k_n} \right) \right). \tag{1}$$

Then there exists an integer $j \geq 0$ depending only on f such that $\Psi(n) \geq d^n$ if $n \leq j$, and $\Psi(n) \geq d^{n-j}$ if $n > j$.

It is easy to show that the above result implies the multiplicative independence of iterates of f.

Corollary 1.2 *Suppose that $f = g/h \in \mathbb{F}(X)$ has degree $d \geq 2$, and is not of the form $aX^{\pm d}$, or $L(X^{p^\ell})$, where $L \in \mathbb{F}(X)$ has degree 1 and ℓ is a positive integer. Then for any integer $n \geq 1$, the iterates $f^{(1)}, \dots, f^{(n)}$ are multiplicatively independent, even up to constants.*

Proof If $(f^{(1)})^{k_1} \dots (f^{(n)})^{k_n} = c, c \in \mathbb{F}$, then Theorem 1.1 ensures $k_n = 0$, as otherwise the degree would be positive. Then we get $k_{n-1} = \dots = k_1 = 0$ recursively. □

In the polynomial case, we also obtain a lower bound on the number of distinct zeros of a multiplicative combination of iterates.

Theorem 1.3 *Suppose $f \in \mathbb{F}[X]$ has degree $d \geq 2$, and has non-vanishing derivative. Let $\mathrm{z}(f)$ denote the number of distinct zeros of f (in an algebraic closure of $\mathbb{F}$), and for an integer n define*

$$Z(n) := \min_{\substack{k_1,\dots,k_n \in \mathbb{Z} \\ k_n \neq 0}} \left(\mathrm{z} \left(\left(f^{(1)} \right)^{k_1} \dots \left(f^{(n)} \right)^{k_n} \right) \right). \tag{2}$$

Let e be the least positive integer k such that $f^{(k)}(0) = 0$, and say that $e = \infty$ if $f^{(k)}(0) \neq 0$ for all $k \geq 1$. Suppose that $f(0) \neq 0$ and $\mathrm{z}(f) > 1$, or that $\mathrm{z}(f) > 2$. Then $Z(n) \geq \gamma(f) d^{n-1} + 1$ if $n \leq e$, and $Z(n) \geq d^{n-e} + 1$ when $n > e$, where

$$\gamma(f) = \begin{cases} \max\{\mathrm{z}(f) - 2, 1\}, & \text{if } \mathbb{F} \text{ has characteristic 0,} \\ 1, & \text{otherwise.} \end{cases}$$

We use Corollary 1.2 in the following extension of the main theorem in [10].

Theorem 1.4 *Let q be a prime power and $n \geq 1$ an integer. Let $g, h \in \mathbb{F}_q[X]$ be coprime with $\deg h, \deg g \leq d = \lceil 2\log_q n \rceil$, and suppose $f = g/h$ satisfies the conditions from Corollary 1.2. Suppose that $\alpha \in \mathbb{F}_{q^n}$ has degree n over $\mathbb{F}_q$ and is a root of $X^m h(X) - g(X)$, where $m = \bar{n} = q^{\lceil \log_q n \rceil}$. Then for*

$$s = \begin{cases} n-1, & f \in \mathbb{F}[X], \\ \lfloor (n-1)/2 \rfloor, & \text{otherwise,} \end{cases}$$

and $t = \lfloor \log_d n \rfloor$, α has order in $\mathbb{F}_{q^n}$ at least

$$\binom{s+t}{t} \prod_{i=0}^{t-1} \frac{1}{d^i}.$$

As an aside we additionally ask, given rational functions $F_1, \ldots, F_n \in \mathbb{F}(X, Y)$ and polynomial $u \in \mathbb{F}[X]$, when $F_1(X, u(X)), \ldots, F_n(X, u(X))$ are multiplicatively dependent. In particular, we find upper bounds on the degree of u such that this is possible, and the number of monic u for which this is the case.

Theorem 1.5 *Suppose $\mathbb{F}$ is a field of characteristic zero, n is a positive integer, and $F_i = G_i/H_i \in \mathbb{F}(X, Y)$ are rational functions for $1 \leq i \leq n$, of respective degrees $d_1 \leq \cdots \leq d_n$ in X and $1 \leq e_1 \leq \cdots \leq e_n$ in Y. For $1 \leq i \neq j \leq n$, define*

$$R_{ij}(X) = \mathrm{Res}_Y(G_i, G_j)\mathrm{Res}_Y(G_i, H_j)\mathrm{Res}_Y(H_i, G_j)\mathrm{Res}_Y(H_i, H_j),$$

where $\mathrm{Res}_Y(P, Q)$ is the resultant of $P, Q \in \mathbb{F}[X, Y]$, considered as polynomials in Y, and set

$$E = \sum_{1 \leq i < n} \sum_{i < j \leq n} \deg R_{ij}.$$

If $R_{ij} \not\equiv 0$ for all $i \neq j$, then there are finitely many monic polynomials $u \in \mathbb{F}[X]$ such that

$$F_1(X, u(X)), \ldots, F_n(X, u(X))$$

are multiplicatively dependent. In particular, such a u has degree not exceeding $E + 2d_n - 1$.

Recalling that the resultant of two polynomials of respective degrees m and n is a polynomial in the coefficients of degree $m + n$, and that each G_i, H_i written as a polynomial in Y, has degree at most e_n, with each coefficient having degree not exceeding d_n. We have that for $i \neq j$, $\deg \mathrm{Res}_Y(G_i, G_j) \leq (e_n + e_n)d_n = 2d_n e_n$, and the same bound holds for all the factors of the polynomial R_{ij} defined above. Thus, counting $\frac{n(n-1)}{2}$ distinct pairs $\{i, j\}$, we obtain $E \leq 4n(n-1)d_n e_n$.

Theorem 1.5 can be applied to the particular scenario of shifting a given set of polynomials by a polynomial u, giving a analogue of results for algebraic numbers from [4] and [7].

Corollary 1.6 *Suppose $\mathbb{F}$ has characteristic zero, n is a positive integer and $f_1, \ldots, f_n \in \mathbb{F}[X]$ are distinct polynomials, not all constant, of respective degrees $d_1 \leq \cdots \leq d_n$ and let*

$$C = d_n \frac{n(n-1)}{2}.$$

Then there are at most $\binom{2C+3d_n-1}{C}$ monic polynomials $u \in \mathbb{F}[X]$ such that

$$f_1 + u, \ldots, f_n + u$$

are multiplicatively dependent. In particular, such a u has degree not exceeding $C + 2d_n - 1$.

The paper is organised with sections corresponding to proofs of the main theorems: In the next section, we collect various results on iterates of rational functions, specifically concerning zeros and poles which are common to different iterates, and the degrees of the numerator and denominator of iterates. We use these results to bound from below the number (counted with multiplicity) of zeros and poles of a given iterate which cannot be found in any of the previous ones. We thus obtain Theorem 1.1. In Sect. 3, we give the proof of a version of [8, Main Theorem], which holds for polynomials over fields of arbitrary characteristic. This is used in conjunction with the general method from Sect. 2 to prove Theorem 1.3. In Sect. 4, we discuss elements of high order in finite fields in a manner analogous to [10,19], but in a slightly more general setting. Finally, in Sect. 5, we use resultants in conjunction with the polynomial ABC-theorem to prove Theorem 1.5.

2 Proof of Theorem 1.1

To prove Theorem 1.1, we need some facts about the composition of rational functions. Let $u = v/w$, $F = G/H \in \mathbb{F}(X)$ be in lowest terms over $\mathbb{F}$, and write

$$u(X) = \frac{v(X)}{w(X)} = \frac{a_l X^l + \cdots + a_s X^s}{b_m X^m + \cdots + b_t X^t}, \ a_l, a_s, b_m, b_t \neq 0,$$

with $\deg u \geq 1$. Let $u \circ F = P/Q$. Recall that the degree of a rational function $f \in \mathbb{F}(X)$, written is lowest terms, is equal to the degree $[\mathbb{F}(X) : \mathbb{F}(f(X))]$, and hence by the product formula for degrees of extensions,

$$\deg u \circ F = (\deg u)(\deg F). \tag{3}$$

Next, we have

$$\frac{P(X)}{Q(X)} = \frac{a_l \left(\frac{G(X)}{H(X)}\right)^l + \cdots + a_s \left(\frac{G(X)}{H(X)}\right)^s}{b_m \left(\frac{G(X)}{H(X)}\right)^m + \cdots + b_t \left(\frac{G(X)}{H(X)}\right)^t}$$
$$= H(X)^{m-l} G(X)^{s-t} \frac{q(X)}{r(X)}, \tag{4}$$

where

$$q(X) = \sum_{i=0}^{l-s} a_{l-i} G(X)^{l-s-i} H(X)^i \text{ and } r(X) = \sum_{i=0}^{m-t} b_{m-i} G(X)^{m-t-i} H(X)^i.$$

Note that a composition of rational functions in lowest terms is itself in lowest terms ([6, Lemma 2.2] is easily extended to our situation). In particular, G, H, q and r are pairwise relatively prime. This means we need not worry about the possibility of factors cancelling after composition. Hence, from (4), whenever $\deg G \neq \deg H$ we have

$$\deg P = \deg H(\deg u - l) + (\deg G)s + \deg F(l - s), \tag{5}$$
$$\deg Q = \deg H(\deg u - m) + (\deg G)t + \deg F(m - t), \tag{6}$$

where P/Q is in lowest terms.

We can use these facts to obtain results about which zeros and poles are common to different iterates of f. It turns out that these relations depend primarily on the earliest iterates of f to have either 0 or a point at infinity as a zero or pole. We hence set the following notation.

Definition 2.1 Write $f^{(k)} = g_k/h_k$ for the k-th iterate of f in lowest terms, and let e be defined as in Theorem 1.3. Further define ϵ, μ and ν to be respectively the smallest positive integers k such that $h_k(0) = 0$, $\deg g_k < \deg h_k$, and $\deg g_k > \deg h_k$. These again take the value ∞ if their respective conditions are not satisfied for any $k \geq 1$.

We first note that there are restrictions on the possible combinations of the values e, ϵ, μ, ν. We make particular use of the next result.

Lemma 2.2 *Suppose $\mu < \nu$ and $\epsilon < \infty$. Then $\mu \leq \epsilon < e < \infty$, and in particular $e = \epsilon + \mu$.*

Proof If $\epsilon < \mu$, then by definition we have $\deg g_\epsilon = \deg h_\epsilon = \deg f^{(\epsilon)} = d^\epsilon$ and $\deg g_{\mu-\epsilon} = \deg h_{\mu-\epsilon} = \deg f^{(\mu-\epsilon)} = d^{\mu-\epsilon}$. Hence, upon setting $u = f^{(\epsilon)}$ and $F = f^{(\mu-\epsilon)}$, (5) gives

$$\deg g_\mu = \deg h_{\mu-\epsilon}(\deg f^{(\epsilon)} - \deg g_\epsilon) + \deg g_{\mu-\epsilon} s + \deg f^{(\mu-\epsilon)}(\deg g_\epsilon - s)$$
$$= d^{\mu-\epsilon}(d^\epsilon - d^\epsilon) + d^{\mu-\epsilon} s + d^{\mu-\epsilon}(d^\epsilon - s) = d^\mu \geq \deg h_\mu.$$

This contradicts the definition of μ, and so we must have $\mu \leq \epsilon$.

Furthermore, if $e < \epsilon$, then $f^{(\epsilon-e)}(0) = f^{(\epsilon-e)}\left(f^{(e)}(0)\right) = f^{(\epsilon)}(0)$, so 0 is a pole of $f^{(\epsilon-e)}$, contradicting the choice of ϵ. Hence we have $\epsilon < e$, and by setting $u = f^{(j)}$, $F = f^{(\epsilon)}$ for a positive integer j, (4) gives that 0 is a zero of $f^{(\epsilon+j)}$ if and only if $\deg g_j < \deg h_j$. Thus $e = \epsilon + \mu$. □

We have the following extension of a result of Gao [10, Lemma 2.2].

Lemma 2.3 *For all integers $k > \ell \geq 1$,*

(i) A zero of $f^{(\ell)}$ is a zero of $f^{(k)}$ if and only if $e < \infty$ and $k \equiv \ell \pmod{e}$.
(ii) A pole of $f^{(\ell)}$ is a pole of $f^{(k)}$ if and only if $\deg g_{k-\ell} > \deg h_{k-\ell}$.
(iii) A pole of $f^{(\ell)}$ is a zero of $f^{(k)}$ if and only if $\deg g_{k-\ell} < \deg h_{k-\ell}$.
(iv) If $\mu < \nu$, then a zero of $f^{(\ell)}$ is a pole of $f^{(k)}$ if and only if $\epsilon < e < \infty$ and $k \equiv \ell - \mu \pmod{e}$.

Proof Let $k > \ell \geq 1$. For part (i), suppose that a zero α of $f^{(\ell)}$ is a zero of $f^{(k)}$. Then $f^{(k)}(\alpha) = f^{(\ell)}(\alpha) = 0$. As $f^{(k)} = f^{(k-\ell)} \circ f^{(\ell)}$, we have

$$f^{(k-\ell)}(0) = f^{(k-\ell)}\left(f^{(\ell)}(\alpha)\right) = f^{(k)}(\alpha) = 0.$$

Thus we must have $e < \infty$, so assume this is the case. If $k \equiv \ell \pmod{e}$, say $k = \ell + je$ where $j \geq 1$, then for any zero β of $f^{(\ell)}$,

$$f^{(k)}(\beta) = f^{(je)}\left(f^{(\ell)}(\beta)\right) = f^{(je)}(0) = 0.$$

Hence any zero of $f^{(\ell)}$ is a zero of $f^{(k)}$. Now, suppose $k \not\equiv \ell \pmod{e}$, say $k = \ell + je + r$ where $u \geq 0$ and $1 \leq r < e$. If $f^{(k)}$ and $f^{(\ell)}$ have a zero in common then, by the above argument, $f^{(je+r)}(0) = f^{(k-\ell)}(0) = 0$. But then

$$f^{(r)}(0) = f^{(r)}(f^{(je)}(0)) = f^{(je+r)}(0) = 0,$$

contradicting the choice of e. Therefore $f^{(k)}$ and $f^{(\ell)}$ have no zero in common when $k \not\equiv \ell \pmod{e}$.

Writing $f^{(k)} = f^{(k-\ell)} \circ f^{(\ell)}$, the second and third parts follow immediately from (4).

Now, suppose that $\mu < \nu$. By definition, we have that $\deg g_k = \deg h_k$ for $1 \leq k < \mu$. Set $u = f^{(j)}$, $F = f^{(\mu)}$, so $f^{(\mu+j)} = u \circ F = P/Q$ as in (4). If $e, \epsilon > j \geq 1$, then in (5) and (6), $s = t = 0$ and so $\deg g_{\mu+j} = \deg h_{\mu+j} = d^{\mu+j}$. We thus note that

$$\deg g_k = \deg h_k = d^k \quad \text{for all } 1 \leq k \neq \mu < \mu + \min\{\epsilon, e\}. \tag{7}$$

Suppose a zero α of $f^{(\ell)}$ is a pole of $f^{(k)}$. Then we have

$$f^{(k-\ell)}(0) = f^{(k-\ell)}\left(f^{(\ell)}(\alpha)\right) = f^{(k)}(\alpha),$$

and so 0 is a pole of $f^{(k-\ell)}$. That is, we indeed have $\epsilon < \infty$. Thus $e = \epsilon + \mu$ by Lemma 2.2. If $k \equiv \ell - \mu \pmod{e}$, say $k = \ell + je - \mu = \ell + (j-1)e + \epsilon$, with $j \geq 1$, then for any zero β of $f^{(\ell)}$,

$$f^{(k)}(0) = f^{(\epsilon)}\left(f^{((j-1)e)}\left(f^{(\ell)}(\beta)\right)\right) = f^{(\epsilon)}\left(f^{((j-1)e)}(0)\right) = f^{(\epsilon)}(0).$$

Thus, any zero of $f^{(\ell)}$ is a pole of $f^{(k)}$. Suppose now that $k = \ell + je + r - \mu$, with $j \geq 1$ and $1 \leq r < e$. If a zero β of $f^{(\ell)}$ is a pole of $f^{(k)}$, then $f^{(k-\ell)}(0) = f^{(k)}(\beta)$, and so 0 is a pole of $f^{(k-\ell)} = f^{((j-1)e+\epsilon+r)}$. Since

$$f^{((j-1)e+\epsilon)}(0) = f^{(\epsilon)}\left(f^{((j-1)e)}(0)\right) = f^{(\epsilon)}(0),$$

0 is also a pole of $f^{((j-1)e+\epsilon)}$ and hence, by part (ii), $\deg g_r > \deg h_r$. This is a contradiction, since from (7) and the definition of μ, $\deg g_k \leq \deg h_k$ for all $1 \leq k < \mu + \min\{\epsilon, e\} = \mu + \epsilon = e$. □

We may also determine facts about the degrees of iterates of f.

Lemma 2.4 *Throughout, if* $\min\{\mu, \nu\} < \infty$, *define*

$$\delta = |\deg g_{\min\{\mu,\nu\}} - \deg h_{\min\{\mu,\nu\}}|,$$

and for a positive integer j, *let* S_j *and* T_j *be respectively the degrees of the lowest order term in* g_j *and* h_j. *Using the notation from Definition 2.1, we have*

(i) If $\nu < \mu$, *then for any integer* $i \geq 1$, $\deg g_{i\nu} = d^{i\nu}$, *and* $\deg h_{i\nu} = d^{i\nu} - \delta^i$. *Moreover,* $\deg g_j = \deg h_j = d^j$ *whenever* $j \not\equiv 0 \pmod{\nu}$.
(ii) If $\mu < \nu$ *and* $\epsilon = e = \infty$, *then* $\deg g_j = \deg h_j = d^j$ *for all* $j \neq \mu$.
(iii) Let $\mu < \nu$, $e < \infty$, *and write* $S_e = S$. *Then* $\deg g_{\mu+j} = d^{\mu+j} - \delta S_j$ *and* $\deg h_{\mu+j} = d^{\mu+j} - \delta T_j$ *for any* $j \geq 1$. *If* $j = ie$ *for some integer* $i \geq 1$, *then* $S_j = S^i$, *and otherwise* $S_j = 0$. *We moreover have the following.*

(a) Suppose $e < \epsilon$. *Then* $T_j = 0$ *for all* $j \geq 1$.
(b) Suppose $\epsilon < e$, *and write* $T_\epsilon = T$. *Then* $S = \delta T$. *If* $j = ie + \epsilon$ *for some integer* $i \geq 0$, *then* $T_j = \delta^i T^{i+1}$, *and otherwise* $T_j = 0$.

Proof Throughout the proof, we will write a given iterate $f^{(k)} = u \circ F = P/Q$, and infer the degrees of its numerator and denominator via the equations (5) and (6).

For the first part, we proceed by induction on i. By definition, $\deg g_j = \deg h_j = \deg f^{(j)} = d^j$ for $1 \leq j < \nu$, and we have $\deg g_\nu = d^\nu$ and $\deg h_\nu = d^\nu - \delta$. This proves the case $i = 1$. Let $i \geq 1$ and suppose that $\deg g_{i\nu} = d^{i\nu}$ and $\deg h_{i\nu} = d^{i\nu} - \delta^i$. For an integer $1 \leq j \leq \nu$, set $u = f^{(j)}$ and $F = f^{(i\nu)}$. If $j < \nu$, we obtain

$$\begin{aligned}\deg g_{i\nu+j} &= \deg h_{i\nu}(\deg f^{(j)} - \deg g_j) + \deg g_{i\nu} S_j + \deg f^{(i\nu)}(\deg g_j - S_j)\\ &= (d^{i\nu} - \delta)(d^j - d^j) + d^{i\nu} S_j + d^{i\nu}(d^j - S_j) = d^{i\nu+j},\end{aligned}$$

and similarly $\deg h_{i\nu+j} = d^{i\nu+j}$. When $j = \nu$, we get

$$\begin{aligned}\deg g_{(i+1)\nu} &= \deg h_{i\nu}(\deg f^{(\nu)} - \deg g_\nu) + \deg g_{i\nu} S_\nu + \deg f^{(i\nu)}(\deg g_\nu - S_\nu)\\ &= (d^{i\nu} - \delta)(d^\nu - d^\nu) + d^{i\nu} S_\nu + d^{i\nu}(d^\nu - S_\nu) = d^{(i+1)\nu},\end{aligned}$$

and

$$\begin{aligned}\deg h_{(i+1)\nu} &= \deg h_{i\nu}(\deg f^{(\nu)} - \deg h_\nu) + \deg g_{i\nu} T_\nu + \deg f^{(i\nu)}(\deg h_\nu - T_\nu)\\ &= (d^{i\nu} - \delta^i)(d^\nu - (d^\nu - \delta)) + d^{i\nu}(d^\nu - \delta) = d^{(i+1)\nu} - \delta^{i+1},\end{aligned}$$

as required. The second part follows from (7).

For the third part, setting $u = f^{(j)}$ and $F = f^{(\mu)}$ gives

$$\begin{aligned}\deg g_{j+\mu} &= \deg h_\mu(\deg f^{(j)} - \deg g_j) + \deg g_\mu S_j + \deg f^{(\mu)}(\deg g_j - S_j)\\ &= d^\mu(d^j - \deg g_j) + (d^\mu - \delta)S_j + d^\mu(\deg g_j - S_j) = d^{j+\mu} - \delta S_j\end{aligned}$$

and similarly $\deg h_{j+\mu} = d^{j+\mu} - \delta T_j$. If we put $u = f^{(e)}$, $F = f^{((i-1)e)}$, induction on i with (4) shows that $S_{ie} = S_e^i = S^i$. Also, by Lemma 2.3 (i), if $j \not\equiv 0 \pmod e$, then no zero of $f^{(e)}$ (in particular 0) is a zero of $f^{(j)}$, and so $S_j = 0$.

For the last part of the proof, we make use of Lemma 2.2. If $e < \epsilon$, we must have $\epsilon = \infty$, and so $T_j = 0$ for all j, proving (a). On the other hand, if $\epsilon < e$, then $e = \epsilon + \mu$. Set $u = f^{(\mu)}$ and $F = f^{(\epsilon)}$ so that (4) gives $S = \delta T$, and thus $S_{ie} = \delta^i T^i$. We similarly obtain $T_{ie+\epsilon} = \delta^i T^{i+1}$. Finally, if j is not equal to $ie + \epsilon$ for any integer $i \geq 0$, then $j \not\equiv \epsilon = e - \mu \pmod e$. Thus, by Lemma 2.3 (iv), no zero of $f^{(e)}$ (in particular 0) is a pole of $f^{(j)}$, and so $T_j = 0$, proving (b). □

Corollary 2.5 *Suppose $\mu < \nu$ and $\epsilon < e < \infty$. Then for a positive integer n,* $\deg g_n < \deg h_n$ *if and only if $n \geq \mu$ and $n - \mu \equiv 0 \pmod e$, and* $\deg g_n > \deg h_n$ *if and only if $n \geq \mu + \epsilon$ and $n - \mu \equiv \epsilon \pmod e$.*

Proof By definition, we have $\deg g_n = \deg h_n$ for $n < \mu$, and $\deg g_n < \deg h_n$ for $n = \mu$. Suppose $n > \mu$, and write $n = \mu + j$, so $j = n - \mu$. Then from Lemma 2.4 (iii), $\deg g_n = d^n - \delta S_j$ and $\deg h_n = d^n - \delta T_j$. Hence $\deg g_n < \deg h_n$ if and only if $S_j > 0$, which occurs precisely when $j = n - \mu \equiv 0 \pmod e$ by Lemma 2.4 (iii).

On the other hand, $\deg g_n > \deg h_n$ if and only if $T_j > 0$, and this happens precisely when $j = n - \mu \equiv \epsilon \pmod e$ by Lemma 2.4 (iii)(b). □

We hence obtain the following result.

Lemma 2.6 *Suppose $\mu < \nu$ and $\epsilon < \infty$ and let $1 \leq \ell < k$. Then*

(i) A zero or pole of $f^{(\ell)}$ is a zero of $f^{(k)}$ if and only if it is a pole of $f^{(k-\mu)}$. In particular, if $k \leq \mu$, then no zero or pole of $f^{(\ell)}$ is a zero of $f^{(k)}$.

(ii) A zero or pole of $f^{(\ell)}$ is a pole of $f^{(k)}$ if and only if it is a zero of $f^{(k-\epsilon)}$. In particular, if $k \leq \epsilon$, then no zero of pole of $f^{(\ell)}$ is a pole of $f^{(k)}$.

Proof Recall that since $\mu < \nu$ and $\epsilon < \infty$, we have $\epsilon < e < \infty$ and $e = \epsilon + \mu$ by Lemma 2.2.

For the first part, by Lemma 2.3 (i) we have that a zero of $f^{(\ell)}$ is a zero of $f^{(k)}$ if and only if $k \equiv \ell \pmod{e}$ (note that since $\ell < k$, this implies $k > \ell + e \geq 1 + \epsilon + \mu > \mu$). Then, by Lemma 2.3 (iv), a zero of $f^{(\ell)}$ is a pole of $f^{(k-\mu)}$ if and only if $k - \mu \equiv \ell - \mu \pmod{e}$, which is an equivalent condition. From Lemma 2.3 (iii), a pole of $f^{(\ell)}$ is a zero of $f^{(k)}$ if and only if $\deg g_{k-\ell} < \deg h_{k-\ell}$. This occurs precisely when $k - \ell \geq \mu$ (and so $k \geq \mu$) and $k - \ell - \mu \equiv 0 \pmod{e}$ by Corollary 2.5. On the other hand, a pole of $f^{(\ell)}$ is a pole of $f^{(k-\mu)}$ if and only if $\deg g_{k-\ell-\mu} > \deg h_{k-\ell-\mu}$. By Corollary 2.5, this happens exactly when $k - \mu \equiv \ell \pmod{e}$, which is again equivalent.

For part (ii), by Lemma 2.3 (iv), a zero of $f^{(\ell)}$ is a pole of $f^{(k)}$ if and only if $k \equiv \ell - \mu \pmod{e}$. Since $e = \epsilon + \mu$, this is equivalent to $k - \epsilon \equiv \ell \pmod{e}$, which implies $k > \epsilon$, and is moreover the precise condition for a zero of $f^{(\ell)}$ to be a zero of $f^{(k-\epsilon)}$ by Lemma 2.3 (i). Furthermore, from Lemma 2.3 (ii), a pole of $f^{(\ell)}$ is a pole of $f^{(k)}$ if and only if $\deg g_{k-\ell} > \deg h_{k-\ell}$. By Corollary 2.5, this is equivalent to $k - \ell - \mu \equiv \epsilon \pmod{e}$. Again by Corollary 2.5, this is equivalent to having $\deg g_{k-\ell-\epsilon} < \deg h_{k-\ell-\epsilon}$, which is in turn equivalent to the given pole of $f^{(\ell)}$ being a zero of $f^{(k-\epsilon)}$, by Lemma 2.3 (iii). □

As we remarked in the introduction, in order to prove multiplicative independence for the iterates of f, it is clearly necessary to show that no iterate of f is a monomial, that is, of the form $f(X) = aX^{\pm d}$. We first look to a result of Silverman [21, Theorem 1]. Recall that two rational functions ϕ, ψ are *linearly conjugate* if there exists a rational function u of degree 1 such that $\phi = u^{-1} \circ \psi \circ u$.

Lemma 2.7 *Suppose there exists a positive integer k such that $f^{(k)} \in \mathbb{F}[X]$. Then either $f \in \mathbb{F}[X]$, f is separable and linearly conjugate to $1/X^d$, or f is not separable and $f(X) = L(X^{p^\ell})$ for some $L \in \mathbb{F}(X)$ of degree 1 and integer $\ell \geq 0$.*

Indeed, if no iterate of f is a polynomial, then certainly none can be a monomial. In fact, in the case where f is separable, we show that a rational function has a monomial iterate if and only if it is itself a monomial. This is not true however, when f is not separable. For example, if $\mathbb{F}$ has characteristic 2, then $f(X) = 1 + 1/X^2$ satisfies $f^{(2)}(X) = \frac{1}{X^4+1}$ and $f^{(3)}(X) = X^8$.

Note that in the case of characteristic 0, some cases of the following can actually be viewed as a corollary of the stronger result [24, Theorem 1], which concerns the number of terms (monomials) of composite polynomials. The results of [24] are further extended to rational functions in [9].

Lemma 2.8 *If $f \in \mathbb{F}(X)$ is neither a monomial, nor of the form $L(X^{p^\ell})$ for some integer $\ell \geq 0$ and $L \in \mathbb{F}(X)$ of degree 1, then $f^{(k)}$ is not a monomial for any $k \geq 1$.*

Proof We begin with the case where $f \in \mathbb{F}[X]$ is a polynomial. First suppose $\mathbb{F}$ has zero characteristic. We proceed by induction on k. That is, suppose $\deg f \geq 2$, and

that f is not a monomial. Then the case where $k = 1$ is trivial. If $f^{(k-1)}$ is not a monomial, we can write

$$f(X) = a_1 X^{d_1} + \cdots + a_s X^{d_s};$$
$$s > 1,\ d = d_1 > \cdots > d_s \geq 0,\ a_1, \ldots, a_s \in \mathbb{F} \setminus \{0\},$$

and

$$f^{(k-1)}(X) = b_1 X^{e_1} + \cdots + b_t X^{e_t};$$
$$t > 1,\ d^{k-1} = e_1 > \cdots > e_t \geq 0,\ b_1, \ldots, b_t \in \mathbb{F} \setminus \{0\}.$$

Hence we have the following cases:

If $d_s = 0$, $e_t \neq 0$, we have that

$$\begin{aligned} f^{(k)}(X) &= f(f^{(k-1)}(X)) \\ &= a_1(b_1 X^{e_1} + \cdots + b_t X^{e_t})^{d_1} + \cdots + a_s \end{aligned}$$

has constant term $a_s \neq 0$. Similarly, if $d_s \neq 0$, $e_t = 0$,

$$\begin{aligned} f^{(k)}(X) &= f^{(k-1)}(f(X)) \\ &= b_1(a_1 X^{d_1} + \cdots + a_s X^{d_s})^{e_1} + \cdots + b_t \end{aligned}$$

has constant term $b_t \neq 0$. If $d_s \neq 0$, $e_t \neq 0$, then

$$\begin{aligned} f^{(k)}(X) &= f(f^{(k-1)}(X)) \\ &= a_1(b_1 X^{e_1} + \cdots + b_t X^{e_t})^{d_1} + \cdots + a_s(b_1 X^{e_1} + \cdots + b_t X^{e_t})^{d_s} \end{aligned}$$

has lowest order term $a_s b_t^{d_s} X^{d_s e_t} \neq 0$, since $a_s \neq 0$, $b_t \neq 0$. Finally, when $d_s = e_t = 0$, if $e_2 > 0$, we have

$$\begin{aligned} f^{(k)}(X) &= f(f^{(k-1)}(X)) \\ &= a_1(b_1 X^{e_1} + b_2 X^{e_2} + \cdots + b_t)^{d_1} + \cdots + a_s. \end{aligned}$$

In this case, the term in $X^{(d_1-1)e_1+e_2}$ has coefficient $d_1 a_1 b_1^{d_1-1} b_2 \neq 0$, since we have $a_1, b_1, b_2 \neq 0$, and $\mathbb{F}$ has 0 characteristic. Otherwise, $e_2 = 0$ and

$$\begin{aligned} f^{(k)}(X) &= f^{(k-1)}(f(X)) \\ &= b_1(a_1 X^{d_1} + a_2 X^{d_2} + \cdots + a_s)^{e_1} + b_2. \end{aligned}$$

Similarly, the term in $X^{(e_1-1)d_1+d_2}$ has coefficient $e_1 b_1 a_1^{e_1-1} a_2 \neq 0$. That is, in all cases $f^{(k)}$ is not a monomial, and we are done.

Now, suppose $\mathbb{F}$ has positive characteristic p, and that $f^{(k)}$ is monomial, say of the form cX^{d^k} with $c \in \mathbb{F} \setminus \{0\}$, for some $k > 1$. We can write

$$f(X) = a_1 X^{d_1 p^\ell} + \cdots + a_t X^{d_t p^\ell} + b,$$

where $a_1, \ldots, a_t \in \mathbb{F} \setminus \{0\}$, $b \in \mathbb{F}$, $t \geq 1$, $\ell \geq 0$, $d_1 > \cdots > d_t \geq 1$, and $p \nmid \gcd(d_1, \ldots, d_t)$.

Here, the degree of f is $d = d_1 p^\ell$. Denote $r = p^\ell$ and let

$$\begin{aligned} v(X) &= a_1 X^{d_1} + \cdots + a_t X^{d_t} + b, \\ w_i(X) &= a_1^{r^{-i}} X^{d_1} + \cdots + a_t^{r^{-i}} X^{d_t} + b^{r^{-i}}, \ i \geq 1. \end{aligned}$$

Since r^i is a power of p, we have for any $i \geq 1$

$$(w_i(X))^{r^i} = a_1 X^{d_1 r^i} + \cdots + a_t X^{d_t r^i} + b = v(X^{r^i}).$$

Hence

$$\begin{aligned} f(X) &= v(X^r), \\ f^{(2)}(X) &= v(v(X^r)^r)) = v\left((w_1(X))^{r^2}\right) = (w_2 \circ w_1(X))^{r^2}. \\ &\vdots \\ f^{(k)}(X) &= (w_k \circ w_{k-1} \circ \cdots \circ w_1(X))^{r^k}, \ \ k \geq 1. \end{aligned}$$

Hence we have

$$w_k \circ w_{k-1} \circ \cdots \circ w_1(X) = c_0 X^{d_1^k},$$

where $c_0 = c^{r^{-k}} \neq 0$, since $c \neq 0$. Differentiating then gives

$$\begin{aligned} &w_k'(w_{k-1} \circ \cdots \circ w_1(X)) \cdot w_{k-1}'(w_{k-2} \circ \cdots \circ w_1(X)) \cdots w_2'(w_1(X)) \cdot w_1'(X) \\ &\quad = d_1^k c_0 X^{d_1^k - 1}. \end{aligned} \tag{8}$$

Since $p \nmid \gcd(d_1, \ldots, d_t)$, $w_i' \neq 0$ for all $i \geq 1$. Thus, the polynomial on the left hand side of (8) is not zero. So $p \nmid d_1$, as otherwise the right hand side would be zero. Since $d_1^k c_0 \neq 0$, the Eq. (8) implies that $w_1'(X)$ divides $X^{d_1^k - 1}$. Therefore w_1' is a monomial. Since $p \nmid d_1$, we must have $p \mid d_i$ for $2 \leq i \leq t$. Hence

$$w_i'(X) = d_1 a_1^{-r^i} X^{d_1 - 1}, \ \ i \geq 1,$$

and so $w_2'(w_1(X)) = d_1 a_1^{-r^2} (w_1(X))^{d_1 - 1}$ is also a factor of $X^{d_1^k - 1}$. If $d_1 > 1$, then w_1 is a monomial and hence f must also be a monomial. If $d_1 = 1$, then $d_1 > \cdots > d_t \geq 1$ implies that $t = 1$. Therefore f is a binomial of the form $aX^{p^\ell} + b$.

Now, suppose $f \notin \mathbb{F}[X]$, and that $f^{(k)}$ is a monomial for some $k \geq 1$. Then in particular, some iterate of f is a polynomial.

If f is separable, then by Lemma 2.7, f is linearly conjugate to $1/X^d$. That is, f has the form

$$f(X) = a + \frac{b}{(X-a)^d}, \qquad a, b \in \mathbb{F}.$$

Then $f^{(2)}(X) = a + b^{1-d}(X-a)^{d^2} \in \mathbb{F}[X]$, which is a monomial if and only if $a = 0$, in which case f is a monomial. Suppose $a \neq 0$. Since f is separable, $d \neq p^\ell$ for any $\ell > 0$, and so, since we have already proved the result for polynomials, no iterate of $f^{(2)}$ is a monomial. That is, $f^{(k)}$ is not a monomial for any even $k \geq 2$ unless f is a monomial. Moreover, we have in this case $\nu = 2 < \mu$, so by Lemma 2.4 (i), $\deg g_k = \deg h_k$, and so $f^{(k)}$ is not a monomial, for all odd k.

Finally, if f is not separable, then by Lemma 2.7, $f^{(k)}$ is not a polynomial, and hence is not a monomial, for any $k \geq 1$ unless f is of the form $L(X^{p^\ell})$ for some $L \in \mathbb{F}(X)$ of degree 1. □

We can now prove Theorem 1.1. Recall that we write $f^{(k)} = g_k/h_k$ in lowest terms, and define δ, S_k, and T_k as in Lemma 2.4, again setting $S = S_e$ and $T = T_\epsilon$ where applicable. Now, where $\Psi(n)$ is defined as in (1), noting that $\mathbb{F}(X)$ is a unique factorisation domain, any zeros or poles of $f^{(n)}$ which can not be found in previous iterates will contribute to the value of $\Psi(n)$ counting multiplicity, since $k_n \neq 0$.

We first consider the case where $\nu < \mu$. Then $\deg g_k \geq \deg h_k$ for all k by Lemma 2.4 (i). Hence $\gcd(g_n, h_k) = 1$ for any $k < n$ by Lemma 2.3 (iii). Moreover, if $n \leq e$, then $\gcd(g_n, g_k) = 1$ for any $k < n$, by Lemma 2.3 (i). In this case, we have $\Psi(n) \geq \deg g_n = d^n$. Suppose $e < \infty$ and $n > e$. Then for $k < n$, a zero of $f^{(k)}$ is a zero of $f^{(n)}$ if and only if $k \equiv n \pmod{e}$ by Lemma 2.3. In this case we also have $k \equiv n - e \pmod{e}$, and so such a zero must also be a zero of $f^{(n-e)}$. Write $u = f^{(e)}$ and $F = f^{(n-e)}$, so (4) gives $g_n = g_{n-e}^S q$, where $S > 0$ and $\gcd(q, g_{n-e}) = 1$. Since $f^{(e)}$ is not a monomial by Lemma 2.8, we have $S < d^e$, and so $\Psi(n) \geq \deg q = d^n - Sd^{n-e} \geq d^{n-e}$.

Now, suppose $\mu < \nu$. If $e = \epsilon = \infty$ or $e < \epsilon$, then by Lemma 2.4 (ii) and (iii)(a), $\deg g_j \leq \deg h_j = d^j$ for all $j \geq 1$, and so $\gcd(h_k, h_n) = 1$ for all $1 \leq k < n$ by Lemma 2.3 (ii). Moreover, $\gcd(g_k, h_n) = 1$ for all $1 \leq k < n$ by Lemma 2.3 (iv). Hence $\Psi(n) \geq \deg h_n = d^n$. Suppose $\epsilon < \infty$. Then $\mu < \epsilon < e < \infty$ by Lemma 2.2. Moreover, if $n \leq \epsilon$, $\gcd(g_k, h_n) = \gcd(h_k, h_n) = 1$ by Lemma 2.6 (ii), and thus we again have $\Psi(n) \geq \deg h_n = d^n$. We hence assume that $\mu \leq \epsilon < n < \infty$.

We now split into a further two cases. Firstly, suppose that $\deg g_\mu > 0$, so that $\delta < d^\mu$. Since $e = \mu + \epsilon > \mu$, we do not have $\mu \equiv 0 \pmod{e}$, and so $S_\mu = 0$ by Lemma 2.4 (iii). Hence, where $u = f^{(\mu)}$ and $F = f^{(n-\mu)}$, (4) gives

$$g_n = h_{n-\mu}^\delta q, \tag{9}$$

for some polynomial q relatively prime to $h_{n-\mu}$. From Lemma 2.6 (i), any zero or pole of a previous iterate $f^{(k)}$, $1 \leq k < n$, which is also a zero of $f^{(n)}$, must be a root of

$h_{n-\mu}$. Hence $\Psi(n) \geq \deg q$, and so we aim to bound $\deg q$ from below. If $n = \mu + ie$ for some integer $i \geq 1$, then $n - \mu = \mu + (i-1)e + \epsilon$, and so by Lemma 2.4 (iii)(b),

$$\begin{aligned}\delta \deg h_{n-\mu} + (\deg g_\mu)d^{n-\mu} &= \delta(d^{n-\mu} - \delta^i T^i) + (d^\mu - \delta)d^{n-\mu} \\ &= d^n - \delta^{i+1}T^i = \deg g_n.\end{aligned}$$

Otherwise, $n - \mu \not\equiv 0 \pmod{e}$, and so $\deg g_n \geq \deg h_n$ by Corollary 2.5. That is, $\deg g_n = d^n$, and so

$$\delta \deg h_{n-\mu} + (\deg g_\mu)d^{n-\mu} \leq \delta d^{n-\mu} + (d^\mu - \delta)d^{n-\mu} = d^n = \deg g_n.$$

Hence from (9)

$$\deg q = \deg g_n - \delta \deg h_{n-\mu} \geq (\deg g_\mu)d^{n-\mu} \geq d^{n-\mu},$$

and therefore $\Psi(n) \geq \deg q \geq d^{n-\mu}$.

On the other hand, where $\deg g_\mu = 0$ and correspondingly $\delta = d^\mu$, we set $u = f^{(\epsilon)}$, and $F = f^{(n-\epsilon)}$. If $\epsilon = \mu$, then by definition $\deg g_\epsilon < \deg h_\epsilon$. Otherwise $0 < \epsilon - \mu < \epsilon$, so $\epsilon - \mu \not\equiv 0, \epsilon \pmod{e}$, and thus $\deg g_\epsilon = \deg h_\epsilon$ by Corollary 2.5. Hence, in (4), $f^{(n)} = h_{n-\epsilon}^{m-l} g_{n-\epsilon}^{-T} q/r$, where $m = \deg h_e \geq \deg g_e = l$. That is,

$$h_n = g_{n-\epsilon}^T r, \tag{10}$$

where r is a polynomial relatively prime to $g_{n-\epsilon}$. From Lemma 2.6 (ii), any zero or pole of a previous iterate $f^{(k)}$, $1 \leq k < n$, which is also a pole of $f^{(n)}$, must be a root of $g_{n-\epsilon}$. Hence $\Psi(n) \geq \deg r = \deg h_n - T \deg g_{n-\epsilon}$. Note that $T < d^\epsilon$, as if this were not the case, by Lemma 2.4 (iii) we would have

$$\deg h_{\mu+\epsilon} = d^{\mu+\epsilon} - \delta T = d^{\mu+\epsilon} - d^\mu d^\epsilon = 0,$$

and $S_{\mu+\epsilon} = S_e = \delta T = d^\mu d^\epsilon$, which implies that $f^{(\mu+\epsilon)}$ is a monomial, contradicting Lemma 2.8. In particular, this means that

$$d^n - T d^{n-\epsilon} \geq d^{n-\epsilon}. \tag{11}$$

Hence, if $n = \mu + ie + \epsilon$ for some integer $i \geq 0$, then $n - \epsilon = \mu + ie$, so by Lemma 2.4 (iii), (10) and (11), we have

$$\deg r = d^n - \delta^{i+1}T^{i+1} - T(d^{n-\epsilon} - \delta^{i+1}T^i) = d^n - T d^{n-\epsilon} \geq d^{n-\epsilon}.$$

Otherwise, $n - \mu \not\equiv \epsilon \pmod{e}$, and so $\deg g_n \leq \deg h_n$ by Corollary 2.5. That is, $\deg h_n = d^n$, and so from (10) and (11)

$$\deg r = d^n - T \deg g_{n-\epsilon} \geq d^n - T d^{n-\epsilon} \geq d^{n-\epsilon}.$$

We conclude that $\Psi(n) \geq \deg r \geq d^{n-\epsilon}$, completing the proof. □

3 Proof of Theorem 1.3

Recall the polynomial ABC-theorem (proved first by Stothers [23], then independently by Mason [14] and Silverman [22]).

Lemma 3.1 *Let $\mathbb{F}$ be a field and let $A, B, C \in \mathbb{F}[X]$ be relatively prime polynomials such that $A + B + C = 0$ and not all of A, B and C have vanishing derivative. Then*

$$\max\{\deg A, \deg B, \deg C\} \leq \deg \operatorname{rad}(ABC) - 1,$$

where, for $f \in \mathbb{F}[X]$, $\operatorname{rad}(f)$ is the product of the distinct monic irreducible factors of f.

We use this to obtain a version of part of the main result of [8]. Namely, we give a lower bound for the number of distinct zeros of a composite polynomial.

Lemma 3.2 *Let $f = g \circ h \in \mathbb{F}[X]$, where $g, h \in \mathbb{F}[X]$, h has non-vanishing derivative, and $\mathrm{z}(g) \geq 2$. Then*

$$\mathrm{z}(f) \geq \gamma(g) \deg h + 1,$$

where γ is defined as in Theorem 1.3.

Proof If $\deg h = 1$, then clearly $\mathrm{z}(f) = \mathrm{z}(g)$. Since $\mathrm{z}(g) \geq 2$, we have

$$\gamma(g) \deg h + 1 = \max\{\mathrm{z}(g) - 1, 2\} \leq \mathrm{z}(g) = \mathrm{z}(f),$$

so assume $\deg h \geq 2$.

In the characteristic 0 case, the result is [8, Main Theorem (i)]. When the characteristic is positive, we proceed in much the same vein. Write

$$f(X) = \prod_{i=1}^{n} (X - \alpha_i)^{f_i}, \quad g(X) = \prod_{j=1}^{t} (X - \beta_j)^{k_j},$$

where the α_i and β_j are respectively the distinct roots of f and g in an algebraic closure of f. Then

$$f(X) = g(h(X)) = \prod_{j=1}^{t} (h(X) - \beta_j)^{k_j}.$$

For $\beta_i \neq \beta_j$, the factors $h(X) - \beta_i$ and $h(X) - \beta_j$ have no zeros in common, so $t \leq n$, and there exists a partition of $\{1, \ldots, n\}$ into disjoint subsets $S_{\beta_1}, \ldots, S_{\beta_t}$, such that

$$h(X) - \beta_j = p_j(X) := \prod_{m \in S_{\beta_j}} (X - \alpha_m)^{l_m},$$

with $l_m k_m = f_m$, for every $j = 1, \dots, t$. Since $t = \mathrm{z}(g) > 1$, we can take $1 \le i < j \le t$, and obtain $h(X) = \beta_i + p_i(X) = \beta_j + p_j(X)$. That is,

$$(\beta_i - \beta_j) + p_i + (-p_j) = 0,$$

where the polynomials on the left-hand side are relatively prime, and in particular, since h has non-vanishing derivative, so does p_i. Thus, applying Lemma 3.1, we have

$$\begin{aligned} \max\{\deg(\beta_i - \beta_j), \deg p_i, \deg(-p_j)\} &= \deg h \\ &\le \deg \mathrm{rad}((\beta_j - \beta_i) p_i p_j) - 1 \le n - 1. \end{aligned}$$

Therefore $n = \mathrm{z}(f) \ge \deg h + 1$. □

We now prove Theorem 1.3. Suppose $f \in \mathbb{F}[X]$ has non-vanishing derivative. Then for any positive integer n,

$$\frac{d}{dX} f^{(n)}(X) = f'(f^{(n-1)}(X)) \cdot f'(f^{(n-2)}(X)) \cdots f'(f(X)) \cdot f'(X) \neq 0.$$

We can hence apply Lemma 3.2 to obtain $\mathrm{z}(f^{(n)}) \ge \gamma(f) d^{n-1} + 1$. As in the proof of Theorem 1.1, any zeros of $f^{(n)}$ which cannot be found in previous iterates will contribute to the value of $Z(n)$, but this time without multiplicity. If $n \le e$, then $\gcd(f^{(k)}, f^{(n)}) = 1$ for all $1 \le k < n$ by Lemma 2.3 (i), and so $Z(n) \ge \mathrm{z}(f^{(n)}) \ge \gamma(f) d^{n-1} + 1$. Suppose that $e < n < \infty$, and write

$$f^{(e)}(X) = X^S \phi(X), \quad S \ge 1, \ \phi(0) \neq 0.$$

Note that any zeros of $f^{(n)}$ which are common with a previous iterate belong to $f^{(n-e)}$ by Lemma 2.3 (i). Now,

$$f^{(n)}(X) = f^{(e)}\left(f^{(n-e)}(X)\right) = \left(f^{(n-e)}(X)\right)^S \phi\left(f^{(n-e)}(X)\right).$$

If $e > 1$, then $\mathrm{z}(f^{(e)}) \ge d^{e-1} + 1 > 2$, and otherwise $\mathrm{z}(f^{(e)}) > 2$ by assumption. Hence $\mathrm{z}(\phi) > 1$, and so by Lemma 3.2, $Z(n) \ge \mathrm{z}\left(\phi\left(f^{(n-e)}\right)\right) \ge \gamma(\phi) d^{n-e} + 1 \ge d^{n-e} + 1$. □

4 Proof of Theorem 1.4

If $f \in \mathbb{F}[X]$, this is the main result of [19], so assume otherwise, in which case we define $s = \lfloor (n-1)/2 \rfloor$. Recall the following lower bound from Lambe [12], on the number of solutions to a linear Diophantine inequality:

Lemma 4.1 *Suppose that m and $x_0, \dots, x_{r-1}$ are positive integers such that $\gcd(x_0, \dots, x_{r-1}) = 1$. Then the number of non-negative integer solutions $a_0, \dots, a_{r-1}$ to the inequality*

$$\sum_{i=0}^{r-1} a_i x_i \leq m,$$

is at least

$$\binom{m+r}{r} \prod_{i=0}^{r-1} \frac{1}{x_i},$$

with equality when $x_0 = \cdots = x_{r-1} = 1$.

Now, set $m = \bar{n}$. Since α is a root of $X^m h(X) - g(X)$, we have $\alpha^m = f(\alpha)$. As m is a power of q, applying the Frobenius automorphism iteratively gives

$$\alpha^{m^i} = f^{(i)}(\alpha),\ i \geq 0. \tag{12}$$

Consider the set

$$S = \left\{ \sum_{i=0}^{t-1} a_i m^i : \sum_{i=0}^{t-1} a_i d^i \leq s,\ \ a_i \geq 0 \right\}.$$

Suppose $a \in S$ has two representations $a = \sum_{i=0}^{t-1} a_i m^i = \sum_{i=0}^{t-1} b_i m^i$. For each i,

$$0 \leq a_i, b_i \leq s < n \leq m,$$

so $\sum_{i=0}^{t-1} a_i m^i$ and $\sum_{i=0}^{t-1} b_i m^i$ are both base-m expansions for a. Hence $a_i = b_i$ for each i, and so S has order equal to the number of non-negative integer solutions to the inequality

$$\sum_{i=0}^{t-1} a_i m^i \leq s.$$

Thus, by Lemma 4.1,

$$\#S \geq \binom{s+t}{t} \prod_{i=0}^{t-1} \frac{1}{d^i}.$$

We will show that the powers α^a, with $a \in S$, are distinct in $\mathbb{F}_{q^n}$, so from Lemma 4.1, α has order at least $\#S$.

Suppose that there exist integers a, b in S such that $\alpha^a = \alpha^b$. Writing $a = \sum_{i=0}^{t-1} a_i m^i$ and $b = \sum_{i=0}^{t-1} b_i m^i$, we have

$$\prod_{i=0}^{t-1} \left(\alpha^{m^i}\right)^{a_i} = \prod_{i=0}^{t-1} \left(\alpha^{m^i}\right)^{b_i}.$$

The equation (12) then gives

$$\prod_{i=0}^{t-1}\left(f^{(i)}(\alpha)\right)^{a_i} = \prod_{i=0}^{t-1}\left(f^{(i)}(\alpha)\right)^{b_i}.$$

Let

$$k_1(X) = \prod_{a_i > b_i} g_i(X)^{a_i - b_i} \prod_{a_i < b_i} h_i(X)^{b_i - a_i}$$

and

$$k_2(X) = \prod_{a_i < b_i} g_i(X)^{b_i - a_i} \prod_{a_i > b_i} h_i(X)^{a_i - b_i}.$$

Then $k_1(\alpha) = k_2(\alpha)$. Since α has degree n and k_1 and k_2 have degree at most

$$\sum_{i=0}^{t-1} \max\{a_i, b_i\}\, d^i \le 2s \le n-1,$$

we have $k_1(X) = k_2(X)$. Thus $\prod_{i=0}^{t-1}\left(f^{(i)}(X)\right)^{a_i - b_i} = 1$. Then $a_i - b_i = 0$ for each i by Corollary 1.2, and hence $a = b$. □

In light of Theorem 1.4, we wish to determine whether such a pair (g, h) of suitable polynomials always exists for all n. If this is so, we can construct a reliable algorithm for finding elements of high order in $\mathbb{F}_{q^n}$. Namely, checking $X^{\bar{n}}h(X) - g(X)$ for irreducible factors of degree n, for each appropriate pair $(g, h) \in \mathbb{F}_q[X]^2$. The case where $h(X) = 1$ is considered in [10], where it is reasonably conjectured, but not proved, that for every n, there exists $g \in \mathbb{F}_q[X]$ with $\deg g \le 2\log_q n$, such that $X^{\bar{n}} - g(X)$ has an irreducible factor of degree n.

For our more general situation, we make the following weaker conjecture,

Conjecture 4.2 *Suppose $n \ge 1$, and let T be the set of pairs $(g, h) \in \mathbb{F}_q[X]^2$ of degree not exceeding $d := \lceil 2\log_q n \rceil$ such that $f = g/h$ satisfies the conditions from Corollary 1.2. Then there exists $(g, h) \in T$ such that $X^{\bar{n}}h(X) - g(X)$ has an irreducible factor of degree n.*

To give some evidence for this conjecture, we first obtain a rough lower bound for the order of T. See [2] for the next lemma, regarding the probability that two polynomials in $\mathbb{F}_q[X]$ are relatively prime.

Lemma 4.3 *Let g and h be randomly chosen from the set of polynomials in $\mathbb{F}_q[X]$ of degree a and b respectively, where a and b are not both zero. Then the probability that g and h are relatively prime is $1 - 1/q$.*

Clearly, every pair $(g, h) \in \mathbb{F}_q[X]^2$ with $\deg g = d$, $\deg h = d-1$ and $\gcd(g, h) = 1$ is an element of T. Thus, Lemma 4.3. gives

$$\begin{aligned} \#T &\geq \left(1 - \frac{1}{q}\right) \cdot (q-1)q^d \cdot (q-1)q^{d-1} \\ &\geq \frac{(q-1)^3}{q^2} q^{4\log_q n} = \frac{(q-1)^3}{q^2} n^4. \end{aligned} \tag{13}$$

Now, consider the following result from [10]:

Lemma 4.4 *Let $P_q(m, n)$ be the probability of a random polynomial in $\mathbb{F}_q[X]$ of degree $m \geq n$ having at least one irreducible factor of degree n. Then*

$$P_q(m, n) \sim \frac{1}{n}, \quad as \quad n \to \infty,$$

uniformly for q and $m \geq n$.

If we model $X^{\bar{n}}h(X) - g(X)$ as a random polynomial in $\mathbb{F}_q[X]$ for each $(g, h) \in T$, Lemma 4.4, in conjunction with (13), suggests that for large n, we expect on the order of n^3 pairs $(g, h) \in T$ such that $X^{\bar{n}}h(X) - g(X)$ has an irreducible factor of degree n. Thus it is plausible that at least one such pair exists.

5 Proof of Theorem 1.5

We now restrict the field $\mathbb{F}$ to having characteristic 0. The key tool of this section is Lemma 3.1, and so the results could perhaps be extended to characteristic p, given stronger conditions to ensure that one of the polynomials A, B or C, to which we apply the theorem, has non-vanishing derivative.

We now prove Theorem 1.5. Suppose $F_1(X, u(X)), \ldots, F_n(X, u(X))$ are multiplicatively dependent, and and assume that no proper subset of these is also multiplicatively dependent, as we can remove functions until this is the case. Then every zero and pole of F_i for $1 \leq i \leq n$ must be a zero or pole of F_j for some $j \neq i$. This is because otherwise we would require $k_i = 0$ in the equation

$$\prod_{\ell=1}^{n} F_\ell(X, u(X))^{k_\ell} = 1, \tag{14}$$

and hence the proper subset $\{F_\ell(X, u(X)) : 1 \leq \ell \leq n,\ \ell \neq i\}$ would be multiplicatively dependent. Hence, if α is a zero or pole or $F_i(X, u(X))$, there exists $j \neq i$ such that $F_i(\alpha, Y)$ and $F_j(\alpha, Y)$ have the common zero or pole $u(\alpha)$, giving $R_{ij}(\alpha) = 0$. Thus, any zero or pole of $F_i(X, u(X))$ for $1 \leq i \leq n$ is a zero of $\prod_{1\leq i<j} \prod_{i<j\leq n} R_{ij}$. In particular, since for all $i \neq j$, R_{ij} is not identically zero, we have

$$\deg \operatorname{rad} \prod_{i=1}^{n} G_i(X, u(X)) H_i(X, u(X)) \leq \sum_{1\leq i<j} \sum_{i<j\leq n} \deg R_{ij} = E. \tag{15}$$

Now, for $1 \le i \le n$, write

$$F_i(X,Y) = \frac{G_i(X,Y)}{H_i(X,Y)} = \frac{\sum_{\nu=0}^{e_i} g_{i,\nu}(X)Y^\nu}{\sum_{\nu=0}^{e_i} h_{i,\nu}(X)Y^\nu},$$

and assume, without loss of generality, that g_{i,e_i} is not identically zero (if it is, we can replace G_i with H_i, and g_{i,e_i} with h_{i,e_i} in the following definitions). For $1 \le i < j \le n$, define

$$P(X) = g_{i,e_i}(X)G_j(X,u(X)), \quad Q(X) = g_{j,e_j}(X)u(X)^{e_j-e_i}G_i(X,u(X)),$$

and $D_{ij}(X) = \gcd(P(X), Q(X))$. Then set

$$A(X) = \frac{P(X)}{D_{ij}(X)}, \quad B(X) = -\frac{Q(X)}{D_{ij}(X)}, \quad C(X) = -(A(X)+B(X)).$$

Then A, B, and C are relatively prime polynomials with $A+B+C=0$.

Suppose $\deg u > 2d_n$. By construction, P and Q have the same degree and same leading coefficient, and hence we have $P \mid Q$ if and only if $P = Q$. If $P = Q$, then

$$\begin{aligned} P(X) - Q(X) = &\sum_{\nu=e_j-e_i}^{e_j} (g_{i,e_i}(X)g_{j,\nu}(X) - g_{j,e_j}(X)g_{i,\nu-e_j+e_i}(X))u(X)^\nu \\ &+ \sum_{\nu=0}^{e_j-e_i-1} g_{i,e_i}(X)g_{j,\nu}(X)u(X)^\nu = 0. \end{aligned}$$

Since $\deg u > 2d_n$, the term in $u(X)^\nu$ in the above expression contains monomials in X of degree between $\nu \deg u$ and $\nu \deg u + 2d_n < (\nu+1)\deg u$. Thus there can be no cancellation between these terms, and so

$$g_{i,e_i}(X)g_{j,\nu}(X) - g_{j,e_j}(X)g_{i,\nu-e_j+e_i}(X) = 0, \quad e_j - e_i \le \nu \le e_j,$$

and

$$g_{i,e_i}(X)g_{j,\nu}(X) = 0, \quad 0 \le \nu < e_j - e_i.$$

We conclude that in fact

$$g_{i,e_i}(X)G_j(X,Y) = g_{j,e_j}(X)Y^{e_j-e_i}G_i(X,Y),$$

but $R_{ij} \not\equiv 0$ implies that $\gcd(G_j(X,Y), G_i(X,Y)) = 1$, and so we must have

$$G_j(X,Y) \mid g_{j,e_j}(X)Y^{e_j-e_i}.$$

This is impossible as $G_j(X, Y)$ has degree $e_j > e_j - e_i$ in Y. Therefore $P \nmid Q$, and so $\deg D_{ij} < \deg P$ gives

$$\deg A = \deg P - \deg D_{ij} = \deg g_{i,e_i} + \deg g_{j,e_j} + e_j \deg u - \deg D_{ij} > 0. \quad (16)$$

Thus A has non-vanishing derivative. Moreover, in C, the term in $u(X)^{e_j}$ cancels out, giving

$$\begin{aligned}\deg C \leq\ & (e_j - 1)\deg u \\ & + \max\{\deg g_{i,e_i} + \deg g_{j,e_j-1}, \deg g_{j,e_j} + \deg g_{i,e_i-1}\} - \deg D_{ij}. \quad (17)\end{aligned}$$

Therefore, we have by Lemma 3.1 and (16),

$$\begin{aligned}\deg A &= \deg g_{i,e_i} + \deg g_{j,e_j} + e_j \deg u - \deg D_{ij} \\ &\leq \max\{\deg A, \deg B, \deg C\} \\ &\leq \deg \mathrm{rad}\, ABC - 1 \\ &\leq \deg \mathrm{rad} G_i(X, u(X))G_j(X, u(X)) + \deg g_{i,e_i} + \deg g_{j,e_j} + \deg C - 1.\end{aligned}$$

From (15), $\deg \mathrm{rad} G_i(X, u(X))G_j(X, u(X)) \leq E$, and so (17) gives

$$\begin{aligned}e_j \deg u - \deg D_{ij} \leq\ & E + (e_j - 1)\deg u + \max\{\deg g_{i,e_i} \\ & + \deg g_{j,e_j-1}, \deg g_{j,e_j} + \deg g_{i,e_i-1}\} - \deg D_{ij} - 1\end{aligned}$$

and hence,

$$\begin{aligned}\deg u &\leq E + \max\{\deg g_{i,e_i} + \deg g_{j,e_j-1}, \deg g_{j,e_j} + \deg g_{i,e_i-1}\} - 1 \\ &\leq E + 2d_n - 1.\end{aligned}$$

Therefore, for $1 \leq i \leq n$, $G_i(X, u(X))$ is a product of at most E distinct irreducible factors, with degree not exceeding $e_n(E + 2d_n - 1) + d_n$. If $w_0, \ldots, w_{E-1}$ are the respective multiplicities of said factors, then up to multiplication by a non-zero constant, the number of possibilities for $G_i(X, u(X))$ is at most the number of non-negative integer solutions to the inequality

$$\sum_{j=0}^{E-1} w_j \leq e_n(E + 2d_n - 1) + d_n,$$

which is at most

$$\binom{e_n(E + 2d_n - 1) + E + d_n}{E} \quad (18)$$

from Lemma 4.1. For each such possibility, say

$$G_i(X, u(X)) = \sum_{j=0}^{d_i} \sum_{k=0}^{e_i} a_{jk} X^j u(X)^k = A \prod_{\ell=0}^{E-1} (X - \alpha_\ell)^{b_\ell},$$

if u is monic then A is uniquely determined. Moreover, we have

$$u(X) \mid A \prod_{\ell=0}^{E-1} (X - \alpha_\ell)^{b_\ell} - \sum_{j=0}^{d_i} a_{j0} X^j,$$

so there are finitely many possibilities for monic u.

For corollary 1.6, let $F_i(X, Y) = f_i(X) + Y$, so $G_i(X, Y) = f_i(X) + Y$ and $H_i(X, Y) = 1$. Then

$$R_{ij}(X) = \mathrm{Res}_Y(g_i, G_j) = f_j(X) - f_i(X)$$

has degree at most d_n, and thus

$$E = \sum_{1 \le i < j} \sum_{i < j \le n} \deg R_{ij} \le d_n \frac{n(n-1)}{2} = C$$

The result follows from substituting this into (18), noting that $e_n = 1$ in this case. □

6 Comments

Considering the case $\nu < \mu$ (which encompasses the polynomial case) of Theorem 1.1, and additionally Theorem 1.3, it is of interest to obtain upper bounds for the value e when it is finite. That is, bounds for the period of 0 under iteration of a polynomial or rational function f. This problem is investigated in various contexts in [5,11,15–17,20]. Bounds on the values of the values of ϵ, μ and ν in the rational function case are similarly of interest.

Another problem is to generalise Theorem 1.3 to rational functions. Our approach used for the polynomial case can plausibly be extended to the situation where $\nu \le \mu$, mirroring the proof of the relevant case in Theorem 1.1, but applying an appropriate version of the main theorem in [8]. Such an extension, however, is not immediate for the case $\mu < \nu$.

Also, note that in the case $\mathbb{F} = \mathbb{C}$, Theorem 1.5 may be able to be generalised to several variables, where $F_i \in \mathbb{C}(X_1, \ldots, X_m, Y)$ and $u \in \mathbb{C}[X_1, \ldots, X_m]$, using an appropriate analogue of Mason's theorem (for example [1, Theorem 2]).

Acknowledgements The author is grateful to Alina Ostafe and Igor Shparlinski for their ideas, comments and encouragement. He would also like to thank the referee for a careful reading and valuable suggestions.

References

1. Bayat, M., Teimoori, H.: A new bound for an extension of Mason's theorem for functions of several variables. Arch. Math. **82**, 230–239 (2004)
2. Benjamin, A., Bennett, C.: The probability of relatively prime polynomials. Math. Mag. **80**, 196–202 (2007)
3. Bombieri, E., Gubler, W.: Heights in Diophantine Geometry, pp. 1–16. Cambridge University Press, Cambridge (2006)
4. Bombieri, E., Masser, D., Zannier, U.: Intersecting a curve with algebraic subgroups of multiplicative groups. Int. Math. Res. Not. **20**, 1119–1140 (1999)
5. Canci, J.K.: Finite orbits for rational functions. Indag. Math. **18**(2), 203–214 (2007)
6. Carter, S.: Rational function decomposition of polynomials. RHUMJ **13**(2), 54–62 (2012)
7. Dubickas, A., Sha, M.: Multiplicative dependence of the translations of algebraic numbers. Revista Matematica Iberoamericana **34**, 1789–1808
8. Fuchs, C., Pethö, A.: On composite rational functions having a bounded number of zeros and poles. Proc. Am. Math. Soc. **139**, 31–38 (2011)
9. Fuchs, C., Zannier, U.: Composite rational functions expressible with few terms. J. Eur. Math. Soc. **14**, 175–208 (2010)
10. Gao, S.: Elements of provable high order in finite fields. Proc. Am. Math. Soc. **127**(6), 1615–1623 (1999)
11. Halter-Koch, F., Konečná, P.: Polynomial cycles in finite extension fields. Math. Slovaca **52**(5), 531–535 (2002)
12. Lambe, T.A.: Bounds on the number of feasible solutions to a knapsack problem. SIAM J. Appl. Math. **26**(2), 302–305 (1974)
13. Lidl, R., Niederreiter, H.: Finite Fields. Addison-Wesley, Reading, MA (1983). (Now distributed by Cambridge University Press)
14. Mason, R.C.: Diophantine Equations Over Function Fields, London Mathematical Society Lecture Note Series, vol. 96. Cambridge University Press, Cambridge (1984)
15. Narkiewicz, W.: Polynomial cycles in cubic fields of negative discriminant. Funct. Approx. Comment. Math. **35**, 261–269 (2006)
16. Narkiewicz, W., Marszalek, R.: Finite polynomial orbits in quadratic rings. Ramanujan J. **12**(1), 91–130 (2006)
17. Narkiewicz, W.: Polynomial cycles in certain rings of rationals. J. Theor. Nombres Bordx. **14**(2), 529–552 (2002)
18. Ostafe, A., Sha, M., Shparlinski, I.E., Zannier, U.: On multiplicative dependence of values of rational functions and a generalisation of the Northcott theorem. Michigan Math. J. **68**, 385–407 (2019)
19. Popovych, R.: On elements of high order in general finite fields. Algebra Discrete Math. **18**(18), 295–300 (2014)
20. Pezda, T.: Polynomial cycles in certain local domains. Acta Arith. **66**(1), 11–22 (1994)
21. Silverman, J.H.: Rational fucntions with a polynomial iterate. J. Algebra **180**(54), 102–110 (1996)
22. Silverman, J.H.: The S-unit equation over function fields. Proc. Camb. Philos. Soc. **95**, 3–4 (1984)
23. Stothers, W.W.: Polynomial identities and Hauptmoduln. Q. J. Math. Oxf. **32**(3), 349–370 (1981)
24. Zannier, U.: On the number of terms of a composite polynomial. Acta Arith. **127**(2), 157–167 (2007)